The Analysis of Verbal Behavior 1991, 9, 41-48

Maximization of Reinforcement by Two Autistic Students with Accurate and Inaccurate Instructions

Bobby Newman, Dawn M. Buffington, and Nancy S. Hemmes
Queens College and the Graduate Center, CUNY, and Queens Services for Autistic Citizens, Inc.

The present study examines maximization of reinforcement by two autistic individuals under conditions of no instructions, accurate instructions, and inaccurate instructions. Accuracy of instructions and magnitude of reinforcement for differential responding in a choice paradigm were systematically varied across phases. Subject one maximized reinforcement across all three conditions in seven experimental phases. Subject two maximized across these same seven phases, but also experienced three additional phases. In two of the additional phases, subject two maximized reinforcement. In a ninth phase, when reinforcement was intermittent rather than continuous, he failed to maximize reinforcement. Implications of the results for the controversies surrounding the concept of rule-governed behavior are discussed.

It has been suggested that one of the key deficits of individuals with autism may be their tendency to perseverate on a particular response and, therefore, their failure to sample alternate schedules or sources of reinforcement (Mullins & Rincover, 1985). Failure to maximize is not peculiar to the autistic. It has been suggested that inaccurate instructions may lead normally developed adult humans to fail to maximize, and to become insensitive to existing contingencies for responding (Buskist, Bennett, & Miller, 1981; Catania, 1986; Higgins & Morris, 1984; Lippman & Meyer, 1967; Lowe, 1979; Matthews, Shimoff, Catania, & Sagvolden, 1977; Shimoff, Catania, & Matthews, 1981). Using inaccurate instructions, several researchers (e.g., Kaufman, Baron, & Kopp, 1966; Weiner, 1970) found strong instructional control over responding even during programmed extinction conditions.

With the evolution of a human verbal community, control of operant behavior by contingencies was supplemented by control via verbal descriptions of contingencies (Skinner, 1963). Thus, the human species acquired the ability to have its behavior controlled by verbal descriptions of the relationships between responses and the consequences that reliably follow these responses.

Control of behavior transfers from instructional control to contingency control, as long as the instructions are consistent with the actual contingencies (Skinner, 1989). Rules (descriptions of contingencies) may help the individual come into contact with the naturally-occurring contingencies. For example, accurate instructions regarding schedules of reinforcement may enhance schedule control of human behavior (Perone, Galizio, & Baron, 1988). If instructions do not accurately describe contingencies, however, this relationship may not hold. It has been suggested that instructions reduce sensitivity to contingencies, possibly by substituting rule-governed for contingency-shaped behavior (e.g., Hayes, Brownstein, Haas, & Greenway, 1986). The interaction between rule-governed and contingency-shaped behavior may be difficult to tease apart, since when instructions are accurate, "Behavior under the control of a rule may occur in the same form as schedule-sensitive behavior and yet be purely an instance of rule-following not controlled by the par-

We would like to acknowledge the contribution of Stanley Sham of Queens College, CUNY for the design and construction of the apparatus used in the current experiment. Reprint requests should be sent to Bobby Newman at Queens College, Psychology Dept., Flushing, NY 11365.

ticular programmed consequences" (Hayes et al., 1986, p. 144).

Baron and Galizio (1983) described instruction following as an operant response class, and suggested that inaccurate instructions will continue to control responding only to the extent that they allow the individual to maximize available reinforcement. They assert that the individual's responding will continue to be instruction-based as long as (s)he does not come into contact with a greater source of available reinforcement, or the instruction-based responding does not lead to reinforcer loss. When instruction-based responding did result in the loss of reinforcers, instructional control was lost (Galizio, 1979).

The current experiment investigated maximization of reinforcement with two autistic individuals. The subjects were higher functioning than those used by Mullins and Rincover (1985). The magnitude of pre-established potent reinforcers associated with two response classes, instruction following and pressing red or green, were manipulated. Instructions, which varied in their accuracy, were presented prior to each trial. Consequences for responding were varied and were presented after each trial.

METHOD

Subjects

Two autistic teenagers, independently diagnosed by an outside agency and known hereafter as Ron and Don, served as the subjects for this experiment. Ron and Don both had IQs in the moderately retarded range, and had developed verbal abilities (i.e., tacts and mands were spontaneously and appropriately emitted). Both were students at an afterschool program for the autistic where the experiment was conducted.

Apparatus

The experimental apparatus consisted of two components, a control box which was connected to a box with lights (red and green). The control box had a dial which allowed the subject to choose which light would be illuminated on the light box. In addition, there was a button in the middle of the control box that activated a bulb on the light box when depressed. In order to illuminate a particular bulb on the light box, the control box dial had to be set to that color setting, and the button had to be depressed. The boxes were placed in front of the subject, the control box on the left of the light box. The dimensions of the control box were 10 x 18 cm, and the dimensions of the light box were 13 x 15 cm. The light bulbs were .5 cm in diameter.

Inter-observer agreement

Inter-observer agreement was taken on both independent and dependent measures for approximately 75% of the blocks. Agreement on all measures was 100% throughout all phases.

Design and Procedure

The experiment was carried out over seven phases for Ron and ten phases for Don. Blocks of trials were run independently for the two subjects. Contingencies for two operant responses, turning on a red or a green light and instruction follow-

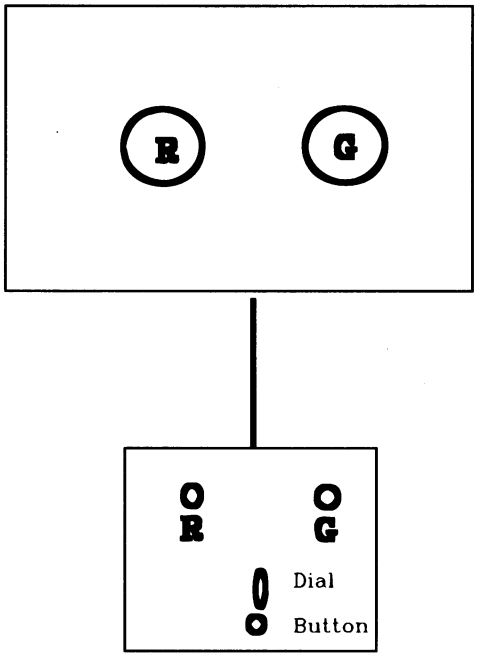

Fig. 1. Experimental apparatus.

ing, were varied across phases (see Table 1 for summary).

Two blocks of ten trials were run each day, with approximately five days per phase. At the beginning of each experimental session, each subject was led into a room in which the apparatus was placed on a circular table. The subject sat at the table, with the apparatus directly in front of him and in easy view. The maximum number of tokens that could be earned in one block (15 tokens in Phases 2 & 7, six tokens in Phase 9, and 20 tokens in the remaining phases) were laid out on the table. The subject was asked to count them. The subject was then asked to "make the red light go on three times, and the green light go on three times." After each response, the appropriate number of tokens (1 or 2 tokens, dependent upon the contingencies specified by the phase) for each light for that session was presented. After the completion of these practice trials, the experimenter turned the control dial to one randomly chosen color setting. After this placement, the day's blocks were begun.

In Phase 1, the subject was instructed to "make the lights go on." Turning on the red light produced two tokens; turning on the green light produced one token. After ten lights had been lit in total, the subject was allowed to trade in his tokens. The trading in of tokens and the consumption of reinforcers separated blocks of trials. In all phases, the maximum number of available tokens (in this phase 20 tokens) was exchangeable for a commodity of the subject's choice (e.g., a can of soda or a package of M & Ms), or 20 minutes in an activity of the student's choosing (e.g., time in the gymnasium or access to a walkman with a favorite music tape). For Ron, when fewer than the maximum number of available tokens were earned, they were exchangeable for an equal number of minutes in an activity of the subject's choosing, except in the final phase. During the seventh phase for Ron, and during the entire experiment for Don, the subject was required to earn the maximum number of available tokens in order to trade in his tokens. Any number of tokens less than the maximum number available was not redeemable (i.e., we switched to an "all or nothing" contingency). This first phase was conducted to verify that the tokens would function as reinforcers, and would

Table 1

Summary of contingencies across experimental phases.

Phase	Trials per	Tokens	*Red Instructed*		*Green Instructed*	
#	Block	Available	Red	Green	Red	Green
1	10	20	2*	1*		
2	10	15	2	0	0	1
3	10	20	2	1	2	1
4	10	20	2	2	2	2
5	10	20	1	2	1	2
6	10	20	2	2	2	2
7	10	15	1	0	0	2
8	10	20	2	0	0	2
9	30	6	p=.2+	p=.1+	p=.2+	p=.1+
10	10	20	2	1	2	1

* No instructions in Phase 1.

+ Intermittent schedules used during this phase, probability of reinforcement presented in table.

control behavior in the absence of instructions. This phase, and all others, was terminated when response patterns appeared stable.

In Phase 2, the experimenter presented an instruction prior to each trial. Instructions were read from a pre-arranged random order of five instructions to turn on the red light, and five instructions to turn on the green light. Instructions were phrased, "Turn on the (color) light and you'll get the most tokens you can," in keeping with Glenn's (1987) conception of rules as descriptions of operant relations, in this case a relation between the required response and a consequence. During this phase, turning on the red light when instructed led to the delivery of two tokens. Turning on the green light when instructed led to the delivery of one token. No tokens were delivered if the instruction was not followed (i.e., the wrong light was turned on). If the subject did not respond after the first instruction, it was repeated a maximum of two more times. If a response did not occur following the second repetition, the next scheduled instruction was delivered. Fifteen tokens were the maximum that could be earned in this phase in any one block of trials. This phase was conducted to demonstrate that the subjects were capable of following instructions, provided the contingencies supported instruction following.

Phase 3 was identical to Phase 2, except there were no differential consequences for compliance vs. noncompliance. Whenever the subject turned on the red light, regardless of instructions, two tokens were delivered. Whenever the subject turned on the green light, regardless of the instructions, one token was delivered. Twenty tokens were the maximum that could be earned in any one block here, and in all subsequent phases except Phase 7 and 9. This phase was conducted to demonstrate that responding would be controlled by contingencies rather than instructions if instruction following did not lead to maximization.

Phase 4 was identical to Phase 3, with the exception that turning on either light was equally reinforced (two tokens for red and two for green). This phase was conducted to examine the response pattern when compliance and noncompliance were not differentially reinforced.

Phase 5 was identical to Phase 4, with the exception that turning on the green light led to the delivery of two tokens, and turning on the red light led to the delivery of one token. This phase was conducted to establish that the response pattern would shift away from the previously established bias towards the red option, if supported by the contingencies.

Phase 6 was identical to Phase 4.

Phase 7 was a replication of Phase 2, with differential consequences for instruction following reinstated. Turning on the green light when so instructed produced two tokens; turning on the red light when so instructed produced one token. Noncompliance with instructions produced no tokens. This phase was conducted to establish that the response class of instruction following could be reinstated when supported by the contingencies.

Both Ron and Don experienced Phases 1 through 7. Only Don, however, experienced Phases 8 through 10. Ron had to drop out of the experiment following the seventh phase due to hospitalization caused by medication difficulties. Ron became inattentive and lethargic as a result of medication changes during Phase 6. Although Ron never failed to respond in any phase, it was during this phase that it became necessary to occasionally repeat instructions. It was this inattentiveness that led us to allow token exchange only after maximization within a given block during Phase 7. It was originally expected that an "all or nothing" contingency would not be able to maintain responding, because it was expected that the subject would never maximize. After seeing that it would be able to maintain responding, the all or nothing contingency was instituted with Don, our second subject.

Phase 8 was identical to Phase 7, except that turning on both red and green led to the presentation of two tokens. This phase was instituted as a control for the fact that 15

tokens was the maximum number of tokens that could be earned in previous accurate instruction phases, as opposed to the 20 that could be earned during the no instruction and inaccurate instruction phases.

In Phase 9, regardless of instructions, both response options led to the delivery of only one token. The red option, however, was reinforced according to a fixed ratio (FR) 5 schedule. The green option was reinforced according to an FR 10 schedule. Due to the intermittency of the reinforcement, each block of this phase contained 30 trials to allow for effective contact with the contingencies.

Phase 10 was identical to Phase 3.

RESULTS

As can be seen in Phase 1 of Figure 2, both subjects maximized with no instructions. The filled circles represent percentage of red chosen. The hollow circles represent percentage of available reinforcers earned. After initially alternating between the two choices, Ron maximized reinforcement on the third block, and again on the fifth through tenth blocks, by responding exclusively on the red light alternative. Don maximized on the first, and then on the fourth through tenth blocks.

As is visible in Phase 2, both subjects also maximized reinforcement with accurate instructions. Despite an initial bias towards the red light, Ron maximized reinforcement on the seventh through tenth blocks of this phase by following instructions on all trials. Don maximized on the second, and then on the fourth through tenth blocks.

In Phase 3, Ron initially failed to maximize reinforcement, complying with all instructions (the first four blocks of this phase). Subsequent performance was variable, finally stabilizing with reinforcement maximization during the ninth block of the phase, and continuing to the end of the phase. Don also failed to maximize early in the phase, following instructions on the first five blocks of the phase and following nine out of ten instructions during the sixth block. Don then maximized twice, resumed instruction following for two blocks, and then maximized on the 11th through the 18th blocks of the phase.

In Phase 4, Ron tended to choose red. Because both choices were equally reinforced regardless of instructions, pressing any combination of buttons resulted in maximization during this phase. During the first two blocks, Ron sampled green on 40% of the trials. During the remaining eight blocks, however, the green light was chosen only once. Don followed instructions for the first six blocks of this phase, but then chose the red alternative during the final six blocks of the phase.

In Phase 5, Ron again maximized reinforcement. The phase began with Ron concentrating responses on the red light, and gradually shifting over to the green light. During the final five blocks of the phase, Ron maximized reinforcement by concentrating responses exclusively on the green light alternative. Don chose the red option on the first two blocks of this phase, but then maximized by concentrating responses on the green alternative for the final eight blocks of the phase.

In Phase 6, responding was variable. In some blocks Ron concentrated responses on one light, in some blocks on the other. Ron tended to stay with whichever light he chose initially, changing the light setting within a block only once. Don began the phase by choosing green for three full blocks. He then gradually changed his responding towards instruction following. Maximization on all blocks was again necessitated by conditions of this phase.

In Phase 7, Ron again maximized reinforcement with accurate instructions. Responding was variable, with Ron often failing to maximize after two or more consecutive blocks of maximization. On the final six blocks of this last phase, however, Ron failed to maximize reinforcement only once. Don began to maximize by the third block of the phase, and continued to maximize by following instructions for the remainder of the phase.

Only Don experienced Phases 8 through 10. During Phase 8, an accurate instruction phase, Don maximized reinforcement in all blocks.

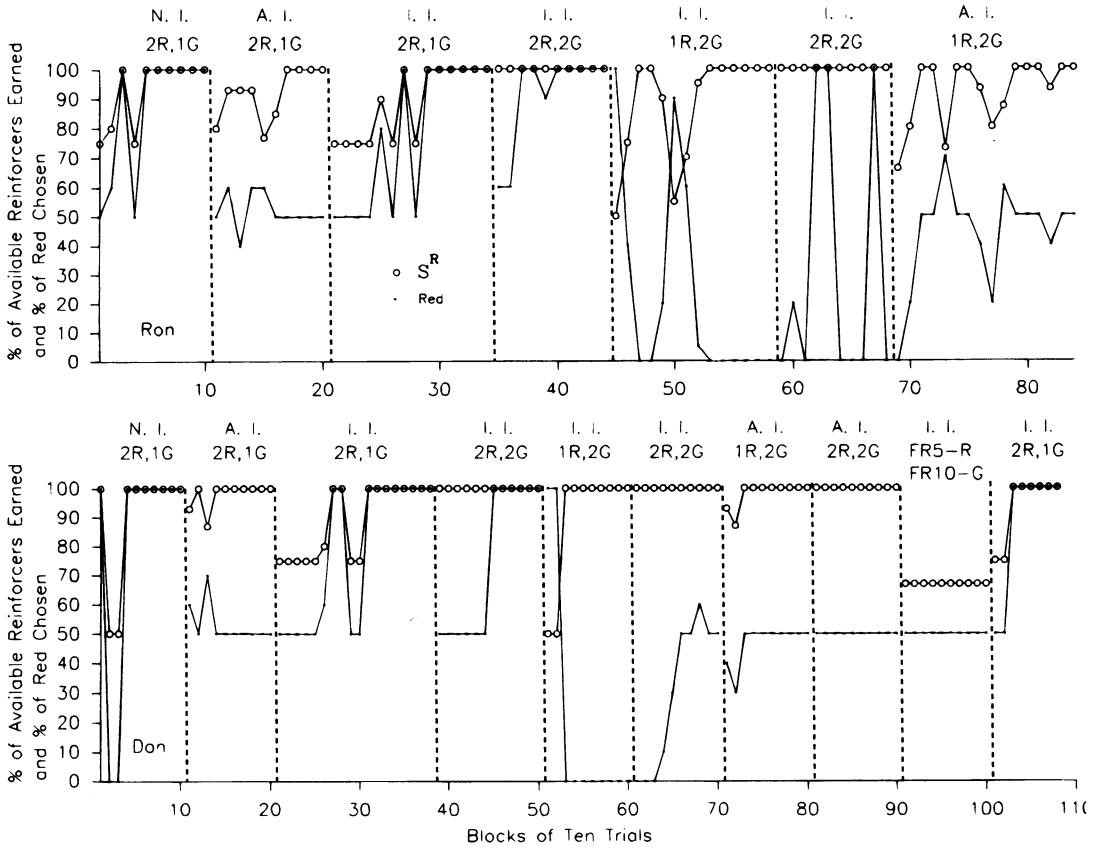

Fig. 2. Percentage of available reinforcers earned, and percentage of red chosen across blocks (N.I. = No Instructions, A.I. = Accurate Instructions, and I.I. = Inaccurate Instructions. Numbers across top refer to the number of tokens delivered after response, R = Red, G = Green).

In Phase 9, Don failed to maximize during any block of the phase, following instructions for a total of 300 trials across ten blocks.

In Phase 10, Don began to maximize with the third block, and continued to maximize for the remainder of the phase.

DISCUSSION

In the current study, reinforcement was maximized under conditions of no instructions, accurate instructions, and inaccurate instructions. The current results run counter to those of Mullins and Rincover (1985). Ron and Don both maximized reinforcement under a number of conditions, some of which required responses contradictory to previous requirements and contrary to the instructions preceding each trial. In view of the current data, it seems that individuals diagnosed with autism can maximize reinforcement, given the appropriate structured environment.

Our results may also have significance for the "insensitivity" issue debated in the rule-governed behavior literature. Little insensitivity to the programmed contingencies was observed in the current study. This is in keeping with the predictions of Baron and Galizio (1983) and the results of DeGrandpre and Buskist (1991).

Insensitivity may have been seen at two points, in the beginning of the third phase (in which there were several full blocks of trials where instructions were followed precisely and reinforcement was thus not maximized), and in the ninth phase. There are two possible explanations for the insensitivity. The first is the response class conception of instruction following, which fits well within the traditional operant framework (Vaughan, 1989) as an acquired function (Reese, 1989). The way in which

responding shifted from instructional to contingency control in Phase 3, only after behavioral variability had allowed for effective contact with the contingencies (e.g., Galizio, 1979), supports this response class conception. Alternatively, we might say that in each phase, rule-governed responding was observed. Responding might have been governed by internally stated rules (e.g., "I have to pick red during this phase in order to maximize", the language hypothesis) rather than the programmed contingencies. Such a conception adds nothing here, however, being entirely post-hoc (Vargas, 1988) and far from parsimonious. Invoking such explanations also leads to potential tautologies.

Phase 9 complicates the issue, however, and raises a methodological issue. Much research in the rule-governed behavior literature has been conducted by giving instructions regarding how to maximize, and then varying schedules of reinforcement without giving any new instructions (Vaughan, 1989). Under these conditions, human subjects have generally failed to show behavior consistent with the new schedule of reinforcement (Hayes, Zettle, & Rosenfarb, 1989; Inesta & Sanchez, 1990). When consequences were continuous, Don showed no insensitivity. He showed insensitivity under the intermittent schedules of reinforcement, but his responding quickly came under contingency control when consequences were once again made continuous.

Insensitivity may therefore be, in part, an artifact of the difficulties humans seem to have with schedules of reinforcement in general, and may or may not be related to the unique linguistic abilities of the human species. We theorize that the intermittent nature of the consequences and their varied order of presentation allowed for adventitious reinforcement of both options, and precluded one option being favored during Phase 9. It may also be the case that instruction following is a "default response," i.e., if you don't know the best way to maximize, your reinforcement history suggests that instruction following is the best course.

Insensitivity may also be due to another methodological issue addressed by our study. Our project differed from many previous studies comparing rule-governed and contingency-shaped behavior in that prior studies have often used reinforcers of dubious strength, e.g., Hayes, et al. (1986), who used points that served as chances in a lottery for two $20 cash prizes to be awarded at the end of a semester. Other researchers have used small monetary reinforcers, e.g., Galizio (1979) who awarded up to a total of $2 for a 50 minute session, and Matthews, Catania, and Sagvolden (1977) who reinforced at a rate of .1 cent per response. Because the subjects in these studies were college students, the potency of such reinforcers as compared to the potential influence of the experiment administrators is questionable (i.e., Baron and Galizio's [1983] argument that such social situations lead to expectations of scrutiny and therefore greater instruction following than would ordinarily be observed). Our study established the potency of our reinforcers before proceeding with the remainder of the experiment. Hayes and Hayes (1989, p. 187) state that "rule-governed behavior is simply behavior controlled by antecedent verbal stimuli." It has further been suggested (Hayes et al., 1989) that the concern with instructions and other forms of verbal stimuli is the defining characteristic of the modern era of operant research. Interestingly, it has also been suggested by Hayes, et al. (1989 p. 215) that research such as the current project, which describes some of the conditions by which subjects will become insensitive to instructions, "trivializes" the literature and "misses the point." If it is true that research regarding rule-governed behavior is the defining characteristic of modern operant research, examining the variables which might contribute to the following of rules would seem to be warranted.

REFERENCES

Baron, A., & Galizio, M. (1983). Instructional control of human operant behavior. *The Psychological Record, 33*, 495-520.

Buskist, W.F., Bennett, R.H., & Miller, H.L. Jr. (1981). Effects of instructional constraints on human fixed-interval performance. *Journal of the Experimental Analysis of Behavior, 35*, 217-225.

Catania, A.C. (1986). On the differences between verbal and nonverbal behavior. *The Analysis of Verbal Behavior, 4,* 2-9.

DeGrandpre, R. J., & Buskist, W. F. (1991). Effects of accuracy of instructions on human behavior: Correspondence with reinforcement contingencies matters. *The Psychological Record, 41,* 371-384.

Galizio, M. (1979). Contingency-shaped and rule-governed behavior: Instructional control of human loss avoidance. *Journal of the Experimental Analysis of Behavior, 31,* 53-70.

Glenn, S.S. (1987). Rules as environmental events. *The Analysis of Verbal Behavior, 5,* 29-32.

Hayes, S.C., Brownstein, A.J., Haas, J.R., & Greenway, D.E. (1986). Instructions, multiple schedules, and extinction: Distinguishing rule-governed from schedule-controlled behavior. *Journal of the Experimental Analysis of Behavior, 46,* 137-147.

Hayes, S.C., & Hayes, L.J. (1989). The verbal action of the listener as a basis for rule-governance. In S.C. Hayes (ed.) *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 153-190). New York: Plenum Press.

Hayes, S.C., Zettle, R.D., & Rosenfarb, I. (1989). Rule-following. In S.C. Hayes (ed.) *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 191-220). New York: Plenum Press.

Higgins, S.T., & Morris, E.K. (1984). Generality of free-operant avoidance conditioning to human behavior. *Psychological Bulletin, 96,* 247-272.

Inesta, E.R., & Sanchez, H. M. (1990). Interaction of contingencies and rule instructions in the performance of human subjects in conditional discrimination. *The Psychological Record, 40,* 565-586.

Kaufman, A., Baron, A., & Kopp, R.E. (1966). Some effects of instructions on human operant behavior. *Psychonomic Monograph Supplements, 1,* 243-250.

Lippman, L.G. & Meyer, M.E. (1967). Fixed-interval performance as related to instructions and subjects' verbalizations of the contingency. *Psychonomic Science, 8,* 135-136.

Lowe, C.F. (1979). Determinants of human operant behavior. In M.D. Zeiler and P. Harzem (Eds.), *Advances in analysis of behaviour: Vol. 1: Reinforcement and the organization of behaviour* (pp. 159-192). New York: John Wiley and Sons.

Matthews, B.A., Shimoff, E., Catania, A.C., & Sagvolden, T. (1977). Uninstructed human responding: Sensitivity to ratio and interval schedules. *Journal of the Experimental Analysis of Behavior, 27,* 453-467.

Mullins, M., & Rincover, A. (1985). Comparing autistic and normal children along the dimensions of reinforcement maximization, stimulus sampling, and responsiveness to extinction. *Journal of Experimental Child Psychology, 40,* 350-374.

Perone, M., Galizio, M., & Baron, A. (1988). The relevance of animal-based principles in the laboratory study of human operant conditioning. In G. Davey and C. Cullen (Eds.) *Human operant conditioning and behavior modification* (pp. 59-86). New York: John Wiley and Sons.

Reese, H.W. (1989). Rules and rule-governance: Cognitive and behavioristic views. In S. C. Hayes (ed.) *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 3-84). New York: Plenum Press.

Shimoff, E., Catania, A.C., & Matthews, B.A. (1981). Uninstructed human responding: Responsivity of low-rate performance to schedule contingencies. *Journal of the Experimental Analysis of Behavior, 36,* 207-220.

Skinner, B.F. (1963). Operant behavior. *American Psychologist, 18,* 503-515.

Skinner, B.F. (1989). *Recent issues in the analysis of behavior.* Toronto: Merrill Publishing Company.

Vargas, E.A. (1988). Event-governed and verbally-governed behavior. *The Analysis of Verbal Behavior, 6,* 11-22.

Vaughan, M. (1989). Rule-governed behavior in behavior analysis. In S.C. Hayes (ed.) *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 97-118). New York: Plenum Press.

Weiner, H. (1970). Instructional control of human operant responding during extinction following fixed-ratio conditioning. *Journal of the Experimental Analysis of Behavior, 13,* 391-394.